Swope, Sheridan

Ex. Search Notes (handwritten)

From: Port, Toby

Sent: Thursday, November 01, 2007 2:07 PM

To: Swope, Sheridan

Subject: FW: BC098632 Sequence Question

Sheridan, does this answer your question? If it doesn't, perhaps you could contact Dr. Temple directly.
Toby

From: Temple, Gary (NIH/NHGRI) [E] [mailto:gtemple@nhgri.nih.gov]
Sent: Thursday, November 01, 2007 1:10 PM
To: Port, Toby
Cc: Temple, Gary (NIH/NHGRI) [E]; Wagner, Lukas (NIH/NLM/NCBI) [E]; Rodriguez, Laura (NIH/NHGRI) [E]; Collins, Francis (NIH/NHGRI) [E]
Subject: RE: BC098632 Sequence Question

Dear Dr Port,

Dr. Collins referred to me your question about the first publication date of BC098632, which is a Mammalian Gene Collection (MGC) cDNA clone. As you noted, the GenBank records show that this cDNA sequence was first published on 1 July 2005, as a submission from the cDNA cloning and sequencing efforts of MGC to provide rat, human, and mouse cDNAs to the research community. The earlier referenced publication, in PNAS 99:16899, 2002, was a summary of the MGC program status, but did not provide individual MGC cDNA sequences, as neither did the Baylor College of Medicine site.

It may be of interest to you that an identical rat cDNA sequence, of 1062 nt, encoding only the full, annotated protein-coding sequence of BC098632 was published 24 July 2003, as GenBank accession no. AY350741, at   http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?
db=Nucleotide&dopt=GenBank&val=37993505. The CDS sequence of this clone is identical to the protein-coding sequence of the longer BC09863, as illustrated at:
http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?
program=blastn&one=37993505&two=BC098632

I hope this gives you the information you need. If you have any further questions, feel free to contact me.

Sincerely,

Gary Temple, M.D.

Program Director, Mammalian Gene Collection (MGC)
National Human Genome Research Institute
National Institutes of Health
5635 Fishers Lane, Suite 4079, MSC 9305 Bethesda, Maryland 20892-9305
Phone: 301-594-5951(Direct); 301-496-7531(Admin)
Fax: 301-480-2770
Email: gtemple@mail.nih.gov

From: Collins, Francis (NIH/NHGRI) [E]
Sent: Thursday, November 01, 2007 10:01 AM
To: Temple, Gary (NIH/NHGRI) [E]
Cc: Rodriguez, Laura (NIH/NHGRI) [E]
Subject: FW: BC098632 Sequence Question

Hi Gary,
    Can you sort this out and provide the requested information to Toby Port?
    Tx, FC
-----Original Message-----
From: Port, Toby [mailto:Toby.Port@USPTO.GOV]
Sent: Wednesday, October 31, 2007 2:58 PM
To: Collins, Francis (NIH/NHGRI) [E]
Cc: Swope, Sheridan
Subject: BC098632 Sequence Question

Dear Dr. Collins,

I am a Librarian at the US Patent & Trademark Office trying to track down when and where Sequence BC098632 (see the attachment) first became publicly available. In the NCBI database, the history reveals that this record became publicly available July 2, 2005.

But when I checked the bibliographic data in the GenEMBL record for this sequence, BC098632 is shown as being in available in an article contributed by you, Francis S. Collins, in the Proceedings of the National Academy of Science in 2002 (see the second attachment), but in fact the complete sequence is not shown in this journal article.

The 2002 journal article states that the rat sequence is NOW available at http://hgsc.bcm.tmc.edu/projects/rat. However, I am not able to find this sequence with or without a date in this database. I already sent this query to the contact address listed at the database, but all we received was a not-too-helpful history of the database and no answer about when the sequence BC098632 first appeared.

Are you able to check the history for this sequence and let me know when it first came online?

Thank you for your help with this request.

Toby Port
US Patent & Trademark Office
STIC-EIC 1600/2900
Remsen 1A59
571-272-2523

<<BC098632.doc>> <<strausberg.pdf>>

Database :   GenEmbl:*
.RGE

RESULT 3
BC098632
LOCUS   BC098632   3039 bp   mRNA   linear   ROD 15-JUL-2006
DEFINITION Rattus norvegicus A kinase (PRKA) anchor protein 7, mRNA (cDNA
            clone MGC:112559 IMAGE:7315112), complete cds.
ACCESSION   BC098632
VERSION   BC098632.1 GI:68534165
KEYWORDS   MGC.
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1 (bases 1 to 3039)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE   Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  PUBMED   12477932
REFERENCE   2 (bases 1 to 3039)
  CONSRTM   NIH MGC Project
  TITLE   Direct Submission
  JOURNAL   Submitted (01-JUL-2005) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
            College of Wisconsin
            cDNA Library Preparation: Express Genomics
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA 94305
            Web site: http://www-shgc.stanford.edu
            Contact: (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found

                    Series: IRAK Plate: 225 Row: d Column: 17.
FEATURES   Location/Qualifiers
     source   1. .3039
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="MGC:112559 IMAGE:7315112"
                     /tissue_type="Brain, rat"
                     /clone_lib="NIH_MGC_254"
                     /lab_host="DH10B"
                     /note="Vector: pExpress1"
     gene   1. .3039
                     /gene="Akap7"
                     /note="synonyms: AKAP18d, Akap15, Akap18"
                     /db_xref="GeneID:361458"
                     /db_xref="RGD:1303071"
     CDS   118. .1179
                     /gene="Akap7"
                     /codon_start=1
                     /product="A kinase (PRKA) anchor protein 7"
                     /protein_id="AAH98632.1"
                     /db_xref="GI:68534166"
                     /db_xref="GeneID:361458"
                     /db_xref="RGD:1303071"
                     /translation="MERPAAGEIDANKCDHLSRGEEGTGDLETSPVGSLADLPFAAVD
                     IQDDCGLPDVPQGNVPQGNPKRSKENRGDRNDHVKKRKKAKKDYQPNYFLSIPITNKK
                     ITAGIKVLQNSILRQDNRLTKAMVGDGSFHITLLVMQLLNEDEVNIGTDALLELKPFV
                     EEILEGKHLTLPFHGIGTFQGQVGFVKLADGDHVSALLEIAETAKRTFQEKGILAGES
                     RTFKPHLTFMKLSKAPMLWKKGVRKIEPGLYEQFIDHRFGEEILYQIDLCSMLKKKQS
                     NGYYHCESSIVIGEKDRKEPEDAELVRLSKRLVENAVLKAVQQYLEETQNKKQPGEGN
                     SVKAEEGDRNGDGSDNNRK"
ORIGIN

   Query Match   100.0%; Score 1062; DB 6; Length 3039;
   Best Local Similarity   100.0%; Pred. No. 3.4e-299;
   Matches 1062; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGGAGCGCCCCGCCGCGGGAGAAATAGATGCCAATAAGTGTGATCATTTATCAAGAGGA 60
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   118 ATGGAGCGCCCCGCCGCGGGAGAAATAGATGCCAATAAGTGTGATCATTTATCAAGAGGA 177

Qy   61 GAGGAAGGGACGGGGGACCTGGAGACCAGCCCTGTAGGTTCTCTGGCAGACCTGCCGTTT 120
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   178 GAGGAAGGGACGGGGGACCTGGAGACCAGCCCTGTAGGTTCTCTGGCAGACCTGCCGTTT 237

Qy   121 GCTGCCGTAGACATTCAAGATGACTGTGGACTCCCTGATGTACCTCAAGGAAATGTACCT 180
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   238 GCTGCCGTAGACATTCAAGATGACTGTGGACTCCCTGATGTACCTCAAGGAAATGTACCT 297

Qy   181 CAAGGAAACCCAAAGAGAAGCAAAGAAAATAGAGGCGACAGGAATGATCACGTGAAGAAG 240
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   298 CAAGGAAACCCAAAGAGAAGCAAAGAAAATAGAGGCGACAGGAATGATCACGTGAAGAAG 357

Qy   241 AGGAAGAAGGCCAAGAAAGATTATCAACCCAACTATTTCCTGTCCATTCCAATCACCAAC 300
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   358 AGGAAGAAGGCCAAGAAAGATTATCAACCCAACTATTTCCTGTCCATTCCAATCACCAAC 417

Qy   301 AAAAAGATTACAGCTGGAATTAAAGTCTTGCAAAATTCGATACTGAGACAGGATAATCGA 360
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db      418 AAAAAGATTACAGCTGGAATTAAAGTCTTGCAAAATTCGATACTGAGACAGGATAATCGA 477

Qy      361 TTGACCAAAGCCATGGTCGGCGACGGCTCCTTTCACATCACCTTGCTAGTGATGCAGCTA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      478 TTGACCAAAGCCATGGTCGGCGACGGCTCCTTTCACATCACCTTGCTAGTGATGCAGCTA 537

Qy      421 TTAAACGAAGATGAAGTAAACATAGGTACCGACGCGCTTTTGGAACTGAAGCCGTTCGTT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      538 TTAAACGAAGATGAAGTAAACATAGGTACCGACGCGCTTTTGGAACTGAAGCCGTTCGTT 597

Qy      481 GAGGAGATCCTTGAGGGGAAGCATCTGACTTTGCCCTTCCACGGGATTGGCACTTTCCAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      598 GAGGAGATCCTTGAGGGGAAGCATCTGACTTTGCCCTTCCACGGGATTGGCACTTTCCAA 657

Qy      541 GGTCAGGTTGGCTTTGTGAAGCTGGCAGACGGAGATCACGTCAGTGCCCTCCTGGAGATA 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      658 GGTCAGGTTGGCTTTGTGAAGCTGGCAGACGGAGATCACGTCAGTGCCCTCCTGGAGATA 717

Qy      601 GCAGAGACTGCAAAAAGGACATTTCAGGAAAAAGGCATCCTGGCTGGAGAAAGCAGAACT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      718 GCAGAGACTGCAAAAAGGACATTTCAGGAAAAAGGCATCCTGGCTGGAGAAAGCAGAACT 777

Qy      661 TTTAAGCCTCACCTGACCTTTATGAAGCTGTCCAAAGCACCAATGCTCTGGAAGAAGGGA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      778 TTTAAGCCTCACCTGACCTTTATGAAGCTGTCCAAAGCACCAATGCTCTGGAAGAAGGGA 837

Qy      721 GTGAGAAAAATAGAGCCTGGATTGTATGAGCAATTTATCGACCACAGATTTGGAGAAGAA 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      838 GTGAGAAAAATAGAGCCTGGATTGTATGAGCAATTTATCGACCACAGATTTGGAGAAGAA 897

Qy      781 ATACTGTACCAAATAGATCTCTGCTCCATGCTGAAGAAAAAACAGAGCAATGGTTATTAC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      898 ATACTGTACCAAATAGATCTCTGCTCCATGCTGAAGAAAAAACAGAGCAATGGTTATTAC 957

Qy      841 CACTGCGAGTCTTCGATCGTGATCGGTGAGAAGGACCGAAAGGAGCCTGAGGATGCTGAA 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      958 CACTGCGAGTCTTCGATCGTGATCGGTGAGAAGGACCGAAAGGAGCCTGAGGATGCTGAA 1017

Qy      901 CTGGTCAGGCTCAGTAAGAGGCTGGTGGAGAACGCCGTGCTCAAGGCTGTCCAGCAGTAC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1018 CTGGTCAGGCTCAGTAAGAGGCTGGTGGAGAACGCCGTGCTCAAGGCTGTCCAGCAGTAC 1077

Qy      961 CTAGAAGAGACACAGAACAAAAAGCAGCCGGGGGAGGGGAACTCCGTCAAAGCTGAGGAG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1078 CTAGAAGAGACACAGAACAAAAAGCAGCCGGGGGAGGGGAACTCCGTCAAAGCTGAGGAG 1137

Qy     1021 GGAGATCGGAATGGCGATGGCAGTGATAACAACCGGAAGTGA 1062
            ||||||||||||||||||||||||||||||||||||||||||
Db     1138 GGAGATCGGAATGGCGATGGCAGTGATAACAACCGGAAGTGA 1179
```

EMBL-EBI | All Databases ▼ | Enter Text Here | ■ Reset Advanced

EBI Databases Database Browsing Dbfetch

EBI Dbfetch

```
ID    BC098632; SV 1; linear; mRNA; STD; ROD; 3039 BP.
XX
AC    BC098632;
XX
DT    04-JUL-2005 (Rel. 84, Created)
DT    16-JUL-2006 (Rel. 88, Last updated, Version 5)
XX
DE    Rattus norvegicus A kinase (PRKA) anchor protein 7, mRNA (cDNA clone
DE    MGC:112559 IMAGE:7315112), complete cds.
XX
KW    MGC.
XX
OS    Rattus norvegicus (Norway rat)
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC    Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea;
OC    Muridae; Murinae; Rattus.
XX
RN    [1]
RP    1-3039
RX    DOI; 10.1073/pnas.242603899
RX    PUBMED; 12477932.
RG    Mammalian Gene Collection Program Team
RA    Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D.,
RA    Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F.,
RA    Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H.,
RA    Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K.,
RA    Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F.,
RA    Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,
RA    Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
RA    Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J.,
RA    Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M.,
RA    Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X.,
RA    Gibbs R.A., Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S.,
RA    Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA    Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C.,
RA    Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I.,
RA    Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;
RT    "Generation and initial analysis of more than 15,000 full-length human and
RT    mouse cDNA sequences";
RL    Proc. Natl. Acad. Sci. U.S.A. 99(26):16899-16903(2002).
XX
RN    [2]
RC    NIH-MGC Project URL: http://mgc.nci.nih.gov
RP    1-3039
RG    NIH MGC Project
RA    ;
RT    ;
RL    Submitted (01-JUL-2005) to the EMBL/GenBank/DDBJ databases.
RL    National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda,
RL    MD 20892-2590, USA
XX
DR    ASTD; TRAN00000020375.
DR    RZPD; IMAGp998N0615368.
XX
```

```
CC    Contact: MGC help desk
CC    Email: cgapbs-r@mail.nih.gov
CC    Tissue Procurement: Drs. Josef Lazar  & Howard Jacob, Medical


CC    College of Wisconsin
CC    cDNA Library Preparation: Express Genomics
CC    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
CC    DNA Sequencing by: Sequencing Group at the Stanford Human Genome
CC    Center, Stanford University School of Medicine, Stanford, CA  94305
CC    Web site:        http://www-shgc.stanford.edu
CC    Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
CC    Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
CC    R. M.
CC    Clone distribution: MGC clone distribution information can be found
CC    through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
CC    Series: IRAK Plate: 225 Row: d Column: 17.
XX
FH    Key            Location/Qualifiers
FH
FT    source         1..3039
FT                   /organism="Rattus norvegicus"
FT                   /lab_host="DH10B"
FT                   /mol_type="mRNA"
FT                   /clone_lib="NIH_MGC_254"
FT                   /clone="MGC:112559 IMAGE:7315112"
FT                   /tissue_type="Brain, rat"
FT                   /note="Vector: pExpress1"
FT                   /db_xref="taxon:10116"
FT                   /db_xref="RZPD:IMAGp998N0615368"
FT    gene           1..3039
FT                   /gene="Akap7"
FT                   /note="synonyms: AKAP18d, Akap15, Akap18"
FT    CDS            118..1179
FT                   /codon_start=1
FT                   /gene="Akap7"
FT                   /product="A kinase (PRKA) anchor protein 7"
FT                   /db_xref="GOA:Q6JP77"
FT                   /db_xref="UniProtKB/TrEMBL:Q6JP77"
FT                   /protein_id="AAH98632.1"
FT                   /translation="MERPAAGEIDANKCDHLSRGEEGTGDLETSPVGSLADLPFAAVDI
FT                   QDDCGLPDVPQGNVPQGNPKRSKENRGDRNDHVKKRKKAKKDYQPNYFLSIPITNKKIT
FT                   AGIKVLQNSILRQDNRLTKAMVGDGSFHITLLVMQLLNEDEVNIGTDALLELKPFVEEI
FT                   LEGKHLTLPFHGIGTFQGQVGFVKLADGDHVSALLEIAETAKRTFQEKGILAGESRTFK
FT                   PHLTFMKLSKAPMLWKKGVRKIEPGLYEQFIDHRFGEEILYQIDLCSMLKKKQSNGYYH
FT                   CESSIVIGEKDRKEPEDAELVRLSKRLVENAVLKAVQQYLEETQNKKQPGEGNSVKAEE
FT                   GDRNGDGSDNNRK"
XX
SQ    Sequence 3039 BP; 886 A; 648 C; 689 G; 816 T; 0 other;
      gtgccgggga tgctgcgact cgcaggctc tgcgcctccg taggcctggc cgccgcggcc        60
      cgcgcccccg ctgccccgcg ggccccgggc ccgcgcctct gccttcgcgc cgcgaccatg       120
      gagcgcccccg ccgcgggaga aatagatgcc aataagtgtg atcatttatc aagaggagag       180
      gaagggacgg gggacctgga gaccagccct gtaggttctc tggcagacct gccgtttgct       240
      gccgtagaca ttcaagatga ctgtggactc cctgatgtac ctcaaggaaa tgtacctcaa       300
      ggaaacccaa agagaagcaa agaaaataga ggcgacagga atgatcacgt gaagaagagg       360
      aagaaggcca agaaagatta tcaacccaac tatttcctgt ccattccaat caccaacaaa       420
      aagattacag ctggaattaa agtcttgcaa aattcgatac tgagacagga taatcgattg       480
      accaaagcca tggtcggcga cggctccttt cacatcacct tgctagtgat gcagctatta       540
      aacgaagatg aagtaaacat aggtaccgac gcgcttttgg aactgaagcc gttcgttgag       600
      gagatccttg aggggaagca tctgactttg cccttccacg ggattggcac tttccaaggt       660
      caggttggct ttgtgaagct ggcagacgga gatcacgtca gtgccctcct ggagatagca       720
      gagactgcaa aaaggacatt tcaggaaaaa ggcatcctgg ctggagaaag cagaactttt       780
      aagcctcacc tgacctttat gaagctgtcc aaagcaccaa tgctctggaa gaagggagtg       840·
      agaaaaatag agcctggatt gtatgagcaa tttatcgacc acagatttgg agaagaaata       900
      ctgtaccaaa tagatctctg ctccatgctg aagaaaaaac agagcaatgg ttattaccac       960
```

```
        tgcgagtctt cgatcgtgat cggtgagaag gaccgaaagg agcctgagga tgctgaactg      1020
        gtcaggctca gtaagaggct ggtggagaac gccgtgctca aggctgtcca gcagtaccta      1080
        gaagagacac agaacaaaaa gcagccgggg gaggggaact ccgtcaaagc tgaggaggga      1140

        gatcggaatg gcgatggcag tgataacaac cggaagtgag agctgaaccc ggtccgctgc      1200
        ccctccgcta agtcgcagac tgactcgcaa tgtgctagtg aagtgtcttg ttcaagccct      1260
        ggagatcacc tagtgattga cgcgattgat gagttcggtt ttgctgcgac acaacagaaa      1320
        agaatggggt gctgggacca gcagaaggaa ttactttaca gaagaacaac acgcacaagg      1380
        gggagccggc acttcgggcc gcctgcccga ctcaaagggc agagggagag gactggtcgc      1440
        cgacagaata ctgttctgcc gtttacattg cttcgatcct ttgactactt tatctgaggc      1500
        caaaacttgc acacagctat caagtgctaa gttcactttg tcactgttga aatgaccatg      1560
        agtatagtga gtccacaatg tttcctgttt gtccccccca tgtgctttta ccacacagtg      1620
        atctttattt acagtaaatt gagttttgtg taaattatat atatttttgg caaatgcaat      1680
        cttttctatg aaatgtgggt aatgttgtaa aggtttttga gccttatttt gataaagtca      1740
        attgccatat ttaatgtcct ccgttgatat ttgtacttta aatgtatcat ataattttcc      1800
        cccttaggca agaaaccagt tggaacccaa gacttaatta acgaagcttt gcaccgagaa      1860
        aggatggagc tgaagtccaa agtgaaacag atcaaagaac ttttgttaaa gcctgagacc      1920
        caggccaaga ttagaaagga gctttttgaa ggaagagttt ttaataatgg tgacccaact      1980
        gatttcaaca tgactttgcc atagctctcc gctgctgtta cgcttgctca catgtttatg      2040
        aactctcagc ccgttaaata gctctttggt gagtaaccaa agtgttctta cggtctctac      2100
        aaagcccaca aaccaacatt tggtaaggaa ctaacaactt cttgccaaag aaaacgtatt      2160
        tttgccttat cgtggtcacc atcatcaccg tcatcatcgt catcaccacc ataaaattga      2220
        gttttagaat gttccttttg gtatcttact cattttatat aaaaacttct taattagctg      2280
        ttgtaagatg ttccatgggt ccttgcagat aatattatat atatatatat atatatacac      2340
        acatatatgt acatatatat acatatatat acacatatac tcacacactt ttaaaaatcc      2400
        tttatagaca aaaacagcaa aacaaaataa aaccaacaac agtattctaa gggtcacctg      2460
        cctcctgttg atgtggtcct gttacttcaa aggaagcatt gtccgggcca gtccagtctc      2520
        aaggtccttt tgctgagcgt ttgagtgctt attgaggatc agcacttgaa cagacattag      2580
        taagcgtaat cgttgtagtc acgggttcag aatgttttat actatctata ttctctcttt      2640
        cattgatgaa gtacagtttg cttttttttt taattttta tttcttcgtg aacagtgttc      2700
        agggttccta tttcctactc tctgaagatg agcccaagcc tgcgttcttc acggtttgag      2760
        tagcttgcac tggttccttt gtaaacgagc attcttgagt gttatttggg tagtcacttt      2820
        aaaattgctg ctactaatag atgatgggga aagaaagtga ttagagatta aatatataat      2880
        catctcacag tccagtttgc tcgtggattt ttggctattt ctttccactg ggtaaatgat      2940
        gcattaattc atgatgtatt cctttatacg tacctacgtt ttcatgcgtc ataataaaag      3000
        tactctttcc tctaaaaaaa aaaaaaaaaa aaaaaaaa                              3039
    //
```